5SSG2059 Spatial Data Science

Is there a correlation between more education and accessibility to public transport in London?

Due: 11 Jan 2021

Student number: 1931393

1986 words

1 Introduction

In 2014, the Hong Kong metro system opened an extension, finally connecting the University of Hong Kong to the rapid transit network (Fung, 2014). The city's top university has been hampered by its relative inaccessibility by transport due to the mountainous terrain of the Island, often requiring inter-modal interchanges. The presence of higher education was clearly a factor in extending coverage, but the debates in transportation geography has focused on more quantitative analysis, such as optimization of coverage area and time travel analysis. There is relatively little attention about the accessibility of nodes. This is a problem because accessibility to public transportation is not just another mathematical variable in models. Health geographers has extensively documented the positive effects of good public transport accessibility and education in the general wellbeing and wealth of society. However, this is still insufficient as transportation and education are treated as independent variables. It is essential to better understand how transportation and education has complemented each other, to inform policy decisions on both transportation and education planning. The aim of this report is to bridge the gap in the relationship between level of education and accessibility to public transport.

2 Literature review

Most research on public transportation within transportation geography are local to one network and limited in scope. Catchment area coverage and travel time analysis are the most popular topics for mathematical analysis (Carlos Garcia-Palomares et al., 2012). Comparatively, there is less work about the accessibility for the network to people. For example, Nettleton et al. (2007) only gave an example of a measure of transportation accessibility in terms of access to healthcare in the UK. O'Sullivan et al. (2000) and Lei et al. (2010) suggests that isochrones (time travel mapping) is the most natural way to assess accessibility. Potential travel demand is often seen as the starting point for analysis (Sun et al., 2018).

Transportation planning in the UK has shifted from scheme-led to accessibility-orientated (Nettleton et al., 2007). Instead of responding to predicted travel demand, the Department for Transport is starting to focus on why people need to travel in the first place. The needs of travellers, including commuting for education, is the starting point for analysis. It is a welcome improvement but more is needed to bridge transportation and education.

Research on how public transport interacts with other components of urban society is the most extensive for health and wellbeing (Pyrialakou et al., 2016). Giles-Corti et al. (2016) argued that public transportation is key to healthier and more sustainable cities. Badland et al. (2014) agrees that both education and transportation are indicators of liveability and wellbeing, while Hogan et al. (2016) found a correlation between happiness and transport accessibility for younger people specifically. Melis et al. (2015) provided further evidence from Italy that good public transportation can reduce risk of depression by increasing geographical mobility.

Education is treated as just another indicator, independent from transportation. For example, Caragliu et al. (2009) analyzed the EU Urban Audit data and found a positive correlation between urban wealth to the quality of transportation networks, and between urban wealth and the level of education. There seems to be little attention paid to the fact that transportation and education mutually supports each other.

Geographers from South America related transport and education in the context of inequality. Hernandez (2018) found that people with higher income has better access via transportation to higher education and job opportunities, highlighting inequity issues in urban society. Moreno-Monroy et al. (2018) found that the inequality in accessibility to schools by public transport in Sao Paulo is creating spatial inequality, accentuating social and economic inequality. Palm et al. (2020) is a rare study of the direct relationship between transport and education. They found a significant positive relationship between transport accessibility and participation in after-school activities, with a weaker relationship to attending non-local schools, concluding that transportation is significant in student's lives and hence education. This report aims to expand on this relationship.

3 Methodology

The level of education will be represented by the number of people with qualifications at Level 4 or above. The UK classifies qualifications into 9 levels; Level 4 and above is usually a certificate of higher education and above (gov.uk, n.d.). The data is from the 2011 UK population census, at the LSOA-level. The London data is compiled by the Greater London Authority (2011) on London Data Store.

The number of people with higher education qualifications is likely to be reliable as it is sourced from the UK-wide census and is representative of the population. It is continuous which is appropriate for regression analysis. While the number of relatively well-educated people does not necessarily directly measure the overall level of education, it is difficult to quantify level of education into a continuous variable. Furthermore, measures such as "number of people with primary education" are used in datasets like the World Bank.

The accessibility of public transportation will be represented by the Transport for London's (2015b) Public Transport Accessibility Levels (PTALs). The PTAL is calculated, for every evenly-spaced point, based on proximity to public transport, the number of services available, and the frequency of services (ie, waiting time) (Transport for London, 2015a). The average of those points' values is the average PTAL for the LSOA, which is used for the analysis.

The average PTAL is used because it is the most suitable for analysis. The raw number of people in each PTAL is too heavily skewed, for example, only 7% of all LSOAs has people in the highest accessibility level. There is no clear reason to pick any one level for analysis while the average accounts for all levels.

Directly using distance to stations or frequency in minutes is not a feasible option because the TfL has not released such data and computing them is out of scope. Using travel time as a measure is even more limited because of the need to pick a single destination point. It only measures accessibility relative to a central point. Catchment Area analysis also face the same problem. Manually calculating them will be very time consuming because of the factorial time complexity, and can not match the PTAL in data accessibility. The PTAL is the more cost and time efficient measure that is still suitable to measure accessibility.

It is still important to keep in mind that the average PTAL is an abstraction of an ordinal variable. The PTAL is still limited as it does not account for the number of destinations available and ease of interchange. Ivan et al. (2019) found that travellers in Czechia considered the number of connections, the convenience of interchange, avoiding overcrowding, and personal factors such as comfortability. Mavoa et al. (2012) also considered the potential access to destinations in measuring accessibility. The PTAL does not measure any of these variables. These measures also do not show inequality in accessibility within the population, for example between different gender (Kwan, 1999) and for people with disabilities (Church et al., 2003).

4 Findings and discussion

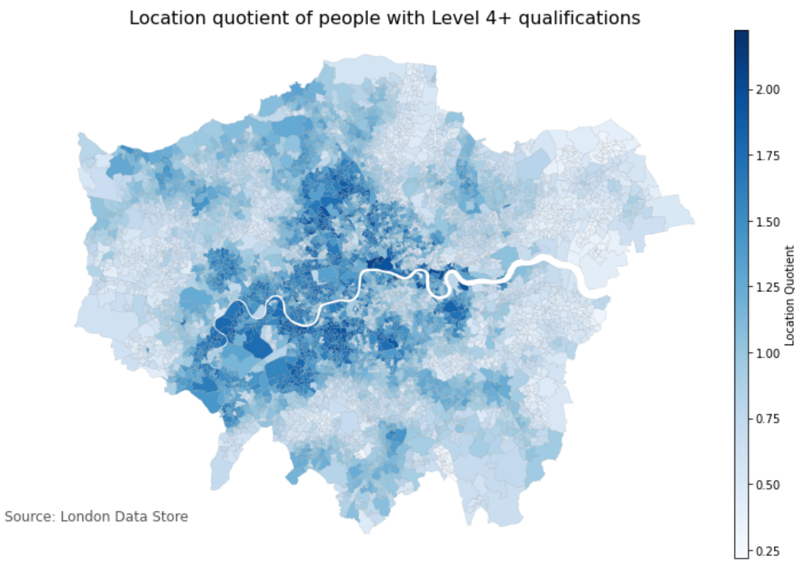


Figure 1: LSOAs coloured by the location quotient of people with qualifications at Level 4 or higher

The data is explored first before attempting analysis, using maps and histograms. Figure 1 is a map of LSOAs colored by the location quotient of people with qualifications at Level

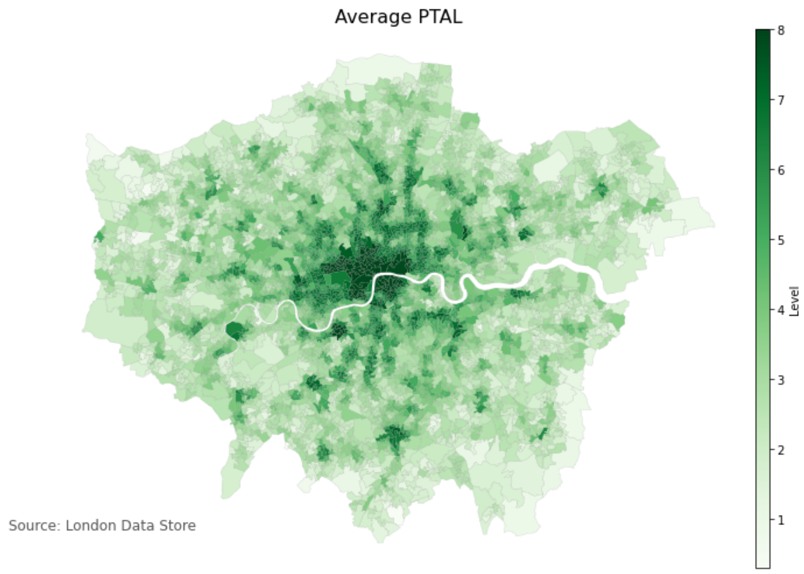


Figure 2: LSOAs coloured by the average Public Transport Accessibility Level

4 or higher. It represents the proportion of well-educated people in each LSOA relative to the entire London. A higher value means there are more highly-qualified people compared to the London-wide proportion. Compare this map with Figure 2, which shows the average PTAL. Both maps show a concentration of high values in central London and lower values in the outskirts, however there is a noticeable concentration of highly-qualified people in south west London.

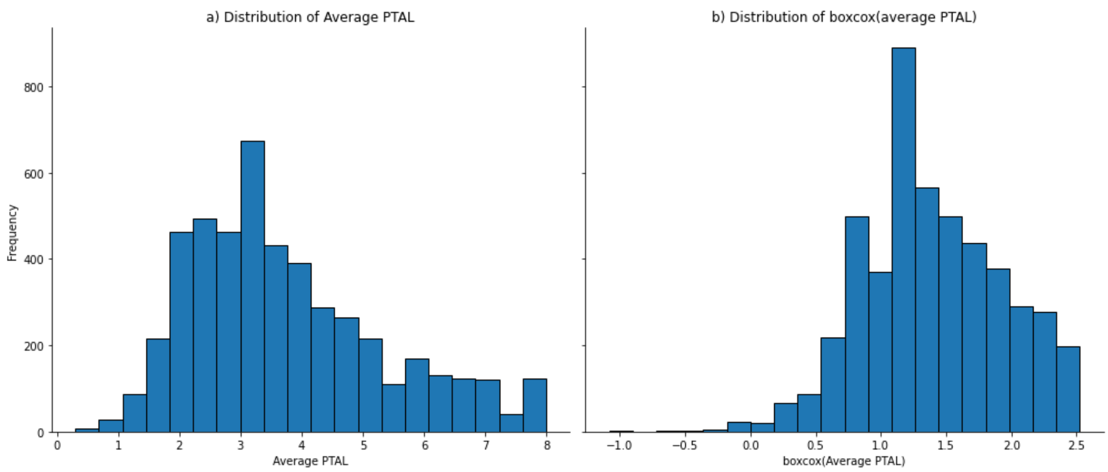


Figure 3: a) A histogram of the average PTAL; b) A histogram of the average PTAL after box-cox transformation

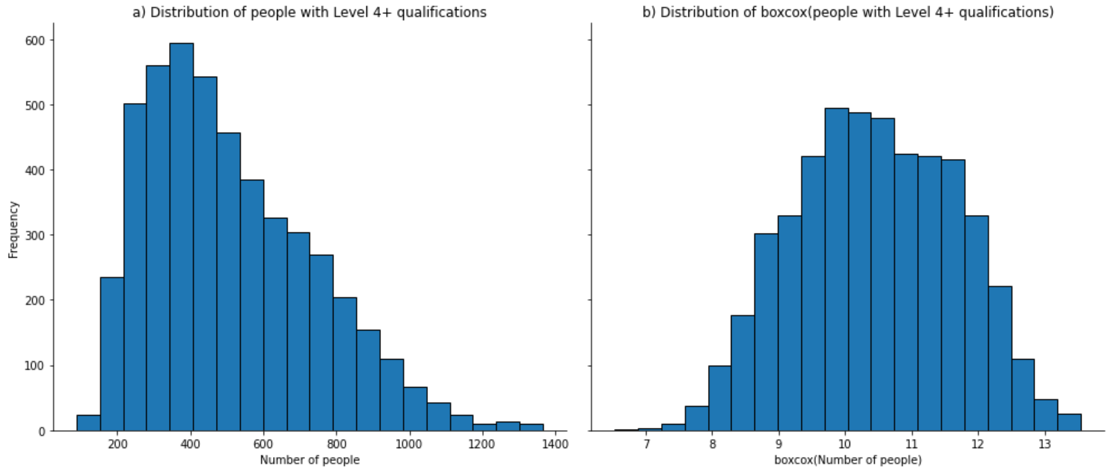


Figure 4: a) A histogram of people with Level 4+ qualifications; b) A histogram of people with Level 4+ qualifications after box-cox transformation

An OLS regression assumes that both variables are normally distributed. Figure 3 shows that the average PTAL is not normally distributed as the skew is 0.794, so it is transformed to reduce the skew. The skew with a log transformation is -0.243 and a box-cox transformation is -0.00249, hence the box-cox transformation is chosen for the regression analysis. The lambda value that was calculated by the transformation is 0.181.

The distribution of people with Level 4+ qualifications has a skew of 0.736 and Figure 4 shows that it looks log-normal, but the skew after a log-transform is -0.163 while a box-cox transform was able to reduce it to -0.0163, hence the box-cox transform will be used. The lambda value is 0.159.

The OLS linear regression is plotted in Figure 5. The Pearson R-squared value is 0.202 and the Spearman rho is 0.457 with a p-value of 2.57×10^{-248}. The gradient is 1.011 with a 95% confidence interval that it is between 0.954 and 1.068. It represents the effect size, which indicates that for a unit increase in box-cox(average PTAL), we expect around a unit increase in box-cox(number of people with Level 4+ qualifications).

The OLS regression needs to satisfy the assumptions made. Figure 6a shows the distribution of standardized residuals. OLS assumes that the residuals are normally distributed, which appears to be true in this plot. Figure 6b is a plot of the standardized residuals against predicted values. OLS assumes that the residuals are normally distributed around zero (the red line), are independent of each other (no correlation visible), and variance is constant. All of those assumptions are satisfied reasonably. Finally, Figure 6c shows a Quantile-Quantile plot of the standardized residuals with a red line showing the ideal normal distribution. While there is deviation on the tails, it looks reasonably normal.

Therefore, it can be concluded that there is a correlation between the level of accessibility

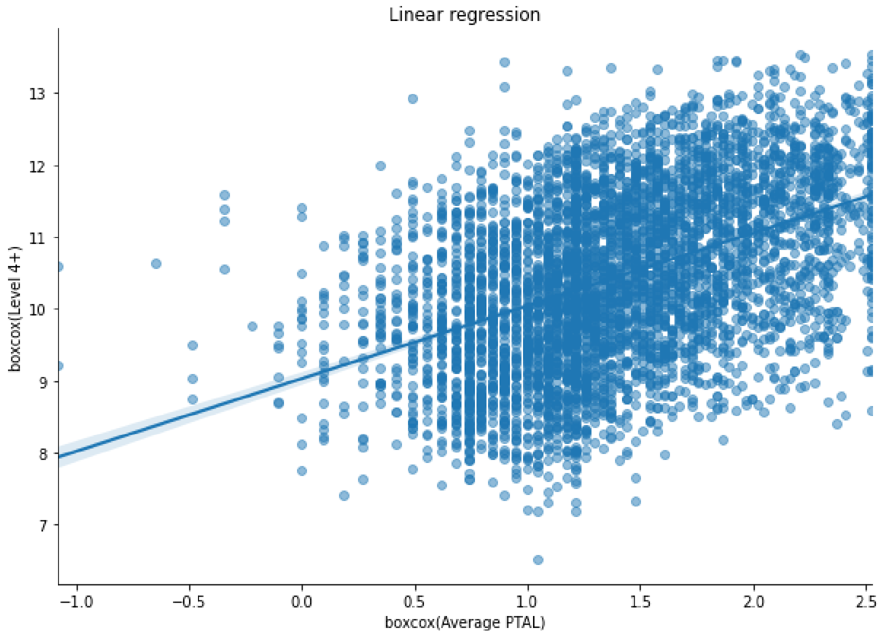
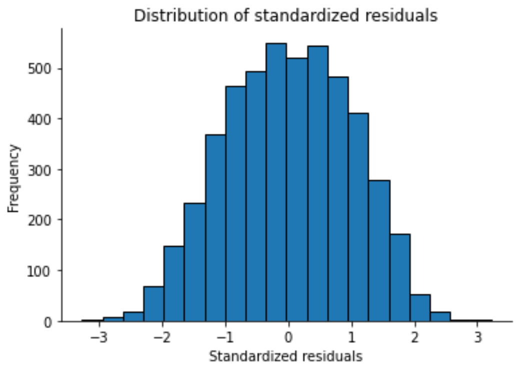
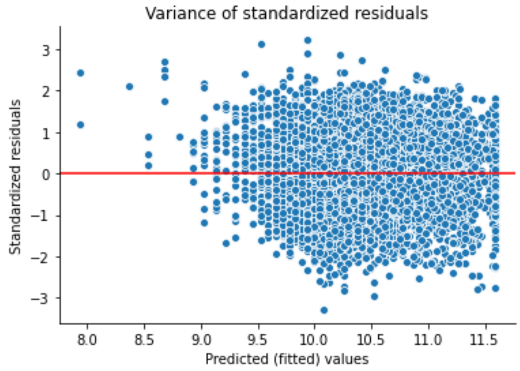


Figure 5: OLS Linear regression between box-cox(Level 4+) and box-cox(Average PTAL)

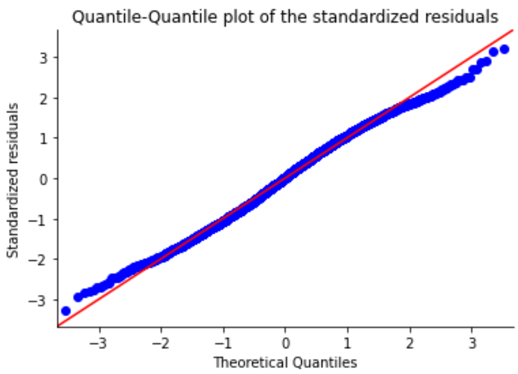
to public transportation and higher qualifications. Specifically there is a linear correlation between both after a box-cox transformation.



(a) Distribution of standardized residuals of the OLS regression



(b) Variance of standardized residuals of the OLS regression



(c) Quantile-Quantile plot of the standardized residuals

Figure 6: Testing the assumptions of OLS

5 Limitations and conclusion

The aim of this report was to investigate the relationship between accessibility to public transportation and level of education. Using data from the TfL and the Census, we found a moderate OLS linear correlation when both variables are transformed with a box-cox transformation. Some remnants of the ordinal data is visible in the scatter plot and the Spearman rho is middling, suggesting a correlation of moderate strength. The assumptions of OLS were satisfied reasonably well.

This agrees with the literature because the level of education and the accessibility to public transport is correlated with another variable in the same direction. For example, the positive correlation of both with urban wealth implies that education and accessibility are also positively correlated themselves.

The correlation does not necessary indicate a causation, because more urban areas tend to be well connected by public transportation because of the high concentration of important services such as government, finance, and higher education. The observed correlation could simply be confirming that higher education institutes choose to locate their campus in areas with good public transport to attract students. Yet, regardless of whether there is a causation or not, it does not affect the policy recommendation that can be made: access to public transportation should be improved to increase the level of education, because being in a well-connected area is a pull factor for individuals in pursuing higher education.

The research should be repeated with other levels of qualifications. The changing effect size for each of those regressions should also confirm that access to public transport is important in improving level of education.

References

Badland, H. et al. (2014). “Urban liveability: Emerging lessons from Australia for exploring the potential for indicators to measure the social determinants of health”. *SOCIAL SCIENCE & MEDICINE* 111, 64–73.

Caragliu, A., C. del Bo, and P. Nijkamp (2009). “Smart cities in Europe”. *CERS 2009 - 3RD CENTRAL EUROPEAN CONFERENCE IN REGIONAL SCIENCE, INTERNATIONAL CONFERENCE PROCEEDINGS - YOUNG SCIENTISTS ARTICLES*. Ed. by Bucek, M and Capello, R and Hudec, O and Nijkamp, P, 45–59.

Carlos Garcia-Palomares, J., J. Gutierrez, and M. Latorre (2012). “Optimizing the location of stations in bike-sharing programs: A GIS approach”. *APPLIED GEOGRAPHY* 35.1-2, 235–246.

Church, R. L. and J. R. Marston (2003). “Measuring Accessibility for People with a Disability”. *Geographical Analysis* 35.1, pp. 83–96.

Fung, F. W. Y. (2014). *Hong Kong MTR’s West Island Line opens on time*. URL: `scmp.com/news/hong-kong/article/1669496/island-line-kennedy-town-extension-opens-time`.

Giles-Corti, B. et al. (2016). “City planning and population health: a global challenge”. *LANCET* 388.10062, 2912–2924.

gov.uk (n.d.). *What qualification levels mean*. URL: `gov.uk/what-different-qualification-levels-mean/list-of-qualification-levels`.

Greater London Authority (2011). *LSOA atlas*. URL: `data.london.gov.uk/dataset/lsoa-atlas`.

Hernandez, D. (2018). “Uneven mobilities, uneven opportunities: Social distribution of public transport accessibility to jobs and education in Montevideo”. *JOURNAL OF TRANSPORT GEOGRAPHY* 67, 119–125.

Hogan, M. J. et al. (2016). “Happiness and health across the lifespan in five major cities: The impact of place and government performance”. *SOCIAL SCIENCE & MEDICINE* 162, 168–176.

Ivan, I. et al. (2019). “Factors Influencing Walking Distance to the Preferred Public Transport Stop in selected urban centres of Czechia”. *GEOSCAPE* 13.1, 16–30.

Kwan, M.-P. (1999). “Gender and Individual Access to Urban Opportunities: A Study Using Space–Time Measures”. *The Professional Geographer* 51.2, pp. 210–227.

Lei, T. L. and R. L. Church (2010). “Mapping transit-based access: integrating GIS, routes and schedules”. *INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE* 24.2, 283–304.

Mavoa, S. et al. (2012). “GIS based destination accessibility via public transit and walking in Auckland, New Zealand”. *JOURNAL OF TRANSPORT GEOGRAPHY* 20.1, 15–22.

Melis, G. et al. (2015). “The Effects of the Urban Built Environment on Mental Health: A Cohort Study in a Large Northern Italian City”. *INTERNATIONAL JOURNAL OF ENVIRONMENTAL RESEARCH AND PUBLIC HEALTH* 12.11, 14898–14915.

Moreno-Monroy, A. I., R. Lovelace, and F. R. Ramos (2018). “Public transport and school location impacts on educational inequalities: Insights from Sao Paulo”. *JOURNAL OF TRANSPORT GEOGRAPHY* 67, 110–118.

Nettleton, M. et al. (2007). “Public Transport Accessibility Map of access to General Practitioners Surgeries in Longbridge, Birmingham, UK”. *JOURNAL OF MAPS*, 64–75.

O’Sullivan, D., A. Morrison, and J. Shearer (2000). “Using desktop GIS for the investigation of accessibility by public transport: an isochrone approach”. *INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE* 14.1, 85–104.

Palm, M. and S. Farber (2020). “The role of public transit in school choice and after-school activity participation among Toronto high school students”. *TRAVEL BEHAVIOUR AND SOCIETY* 19, 219–230.

Pyrialakou, V. D., K. Gkritza, and J. D. Fricker (2016). “Accessibility, mobility, and realized travel behavior: Assessing transport disadvantage from a policy perspective”. *JOURNAL OF TRANSPORT GEOGRAPHY* 51, 252–269.

Sun, C. et al. (2018). “Gated university campus and its implications for socio-spatial inequality: Evidence from students’ accessibility to local public transport”. *HABITAT INTERNATIONAL* 80, 11–27.

Transport for London (2015a). *Connectivity Assessment Guide*.

— (2015b). *Public Transport Accessibility Levels*. URL: data.london.gov.uk/dataset/public-transport-accessibility-levels.